Biometrika (2018), **xx**, x, *pp.* 1–6
Printed in Great Britain

Supplementary material for wild residual bootstrap inference for penalized quantile regression with heteroscedastic errors

By Lan Wang

School of Statistics, University of Minnesota, 224 Church Street South East, Minneapolis, Minnesota 55455, U.S.A.

wangx346@umn.edu

Ingrid Van Keilegom

Research Centre for Operations Research and Business Statistics, KU Leuven, Naamsestraat 69, B-3000 Leuven, Belgium

ingrid.vankeilegom@kuleuven.be

and Adam Maidman

School of Statistics, University of Minnesota, 224 Church Street South East, Minneapolis, Minnesota 55455, U.S.A.

maidm004@umn.edu

Appendix 1

Proofs of Lemma 1, Lemma 2 and Lemma A1

The proofs of Lemmas 1 and 2 combine the ideas in Wu & Liu (2009) and Wang et al. (2012). Section 3.3 of Wu & Liu (2009) considered an extension of the asymptotic theory of penalized quantile regression to the general heteroscedastic error setting but only a sketch of the derivation was provided in their online supplement. We provide a detailed derivation below for completeness.

Proof of Lemma 1. Write $\delta = (\delta_1^T, \delta_2^T)^T$, where $\delta_1 = (\delta_0, \delta_1, \ldots, \delta_q)^T$ and $\delta_2 = (\delta_{q+1}, \ldots, \delta_p)^T$. Write $\widetilde{\delta} = (\widetilde{\delta}_1^T, \widetilde{\delta}_2^T)^T = n^{1/2}(\widetilde{\beta} - \beta_0)$. Then $\widetilde{\delta}$ minimizes $Q_n(\delta)$, where

$$Q_n(\delta) = \sum_{i=1}^{n} \left\{\rho_\tau(\epsilon_i - n^{-1/2}x_i^T\delta) - \rho_\tau(\epsilon_i)\right\} + \lambda_n \sum_{j=1}^{p} w_j\left(|\beta_{0j} + n^{-1/2}\delta_j| - |\beta_{0j}|\right).$$

It follows from Knight (1998) and Koenker (2005) that $\sum_{i=1}^n \left\{\rho_\tau(\epsilon_i - n^{-1/2}x_i^T\delta) - \rho_\tau(\epsilon_i)\right\} = -\delta^T H + \delta^T B_1 \delta/2 + o_p(1)$, where $H \sim N\{0, \tau(1-\tau)B_0\}$. For the penalty term, we consider two cases. (i) For $j = 1, \ldots, q$, $\overline{\beta}_j \to \beta_{0j} \neq 0$ in probability, and $n^{1/2}\big(|\beta_{0j} + \delta_j/n^{1/2}| -$

$|\beta_{0j}|\big) \to \delta_j \mathrm{sign}(\beta_{0j})$. It follows that $\lambda_n w_j\big(|\beta_{0j}+\delta_j/n^{1/2}|-|\beta_{0j}|\big) = (n^{-1/2}\lambda_n)|\overline{\beta}_{0j}|^{-\gamma} n^{1/2}\big(|\beta_{0j}+\delta_j/n^{1/2}|-|\beta_{0j}|\big) \to 0$ as $n^{-1/2}\lambda_n \to 0$. (ii) For $j=q+1,\ldots,p$, $\lambda_n w_j\big(|\beta_{0j}+\delta_j/n^{1/2}|-|\beta_{0j}|\big) = (n^{(\gamma-1)/2}\lambda_n)(n^{1/2}|\overline{\beta}_j|)^{-\gamma}|\delta_j|$. Since $n^{(\gamma-1)/2}\lambda_n \to \infty$ and $n^{1/2}|\overline{\beta}_j| = O_p(1)$, the limit of $\lambda_n w_j\big(|\beta_{0j}+\delta_j/n^{1/2}|-|\beta_{0j}|\big)$ is zero if $\delta_j=0$ and is ∞ if $\delta_j \neq 0$. Hence

$$Q_n(\delta) \to Q(\delta) = \begin{cases} -\delta^T H + \delta^T B_1 \delta/2, & \delta_{q+1}=\cdots=\delta_p=0,\\ +\infty, & \text{otherwise,}\end{cases}$$

in probability. Note that $Q_n(\delta)$ is convex in δ and its limit $Q(\delta)$ has a unique minimum $(D_1^{-1}W, 0_{p-q}^T)^T$, where $W \sim N\{0, \tau(1-\tau)D_0\}$, $D_0 = \lim_{n\to\infty} n^{-1}\sum_{i=1}^n x_{iA}x_{iA}^T$ and $D_1 = \lim_{n\to\infty} n^{-1}\sum_{i=1}^n f_i(0)x_{iA}x_{iA}^T$. It follows from the epi-convergence theory, see Geyer (*On the asymptotics of convex stochastic optimization*, technical report, 1996) and Knight (*Epi-convergence in distribution and stochastic equi-semicontinuity*, technical report, 1999), that $\widetilde{\delta} \to \arg\min_\delta Q(\delta)$ in distribution. Hence $\widetilde{\delta}_1 \to D_1^{-1}W \sim N\{0, \tau(1-\tau)D_1^{-1}D_0D_1^{-1}\}$ in distribution and $\widetilde{\delta}_2 \to 0$ in distribution. This proves (ii).

Note that the above asymptotic normality result suggests that $\mathrm{pr}(j \in \widetilde{A}) \to 1$ for $j=1,\ldots,q$. To prove (i), it remains to show $\mathrm{pr}(j\in\widetilde{A}) \to 0$ for $j=q+1,\ldots,p$. For a given $j \in \{q+1,\ldots,p\}$, let

$$\xi_j(\delta) = -\tau n^{-1/2}\sum_{i=1}^n x_{ij} I(\epsilon_i - n^{-1/2}x_i^T\delta > 0)$$
$$+(1-\tau)n^{-1/2}\sum_{i=1}^n x_{ij} I(\epsilon_i - n^{-1/2}x_i^T\delta < 0) - n^{-1/2}\sum_{i=1}^n x_{ij}v_i + \lambda_n n^{-1/2} w_j \mathrm{sign}(\delta_j),$$

where $v_i = 0$ if $\epsilon_i - n^{-1/2}x_i^T\delta \neq 0$ and $v_i \in [\tau-1,\tau]$ otherwise. By the KKT optimality conditions (Boyd & Vandenberghe, 2004), if $j \in \widetilde{A}$, then there must exist some v_i^* such that $v_i^*=0$ if $\epsilon_i - x_i^T\widetilde{\delta}/n^{1/2} \neq 0$ and $v_i^* \in [\tau-1,\tau]$ otherwise, such that for $\xi_j(\widetilde{\delta})$ with $v_i = v_i^*$, $\xi_j(\widetilde{\delta})=0$. Hence $\mathrm{pr}(j\in\widetilde{A}) \le \mathrm{pr}\{\xi_j(\widetilde{\delta})=0\}$. Note that $\lambda_n n^{-1/2}w_j = (n^{(\gamma-1)/2}\lambda_n)(n^{1/2}|\overline{\beta}_j|)^{-\gamma} \to \infty$ as $n\to\infty$. Furthermore, we have

$$-\tau n^{-1/2}\sum_{i=1}^n x_{ij}I(\epsilon_i - n^{-1/2}x_i^T\widetilde{\delta} > 0) + (1-\tau)n^{-1/2}\sum_{i=1}^n x_{ij}I(\epsilon_i - n^{-1/2}x_i^T\widetilde{\delta} < 0) - n^{-1/2}\sum_{i=1}^n x_{ij}v_i^*$$
$$= n^{-1/2}\sum_{i=1}^n x_{ij}\{I(\epsilon_i - n^{-1/2}x_i^T\widetilde{\delta} \le 0) - \tau\} - n^{-1/2}\sum_{i\in\mathcal{D}} x_{ij}\{v_i^* + (1-\tau)\},$$

where $\mathcal{D} = \{i : \epsilon_i - n^{-1/2}x_i^T\widetilde{\delta} = 0\}$. With probability one the number of elements in $\mathcal{D}$ is finite, following the same argument as in Section 2.2 of Koenker (2005). Therefore, $n^{-1/2}\sum_{i\in\mathcal{D}} x_{ij}\{v_i^* + (1-\tau)\} = O_p(n^{-1/2})$. Similarly as in the proof of Lemma 4.3 of Wang et al. (2012), we can show that for any $\Delta > 0$, as $n \to \infty$,

$$\sup_{||\delta'-\delta||\le\Delta} n^{-1/2}\Big|\sum_{i=1}^n x_{ij}\big\{I(\epsilon_i - n^{-1/2}x_i^T\delta' \le 0) - I(\epsilon_i - n^{-1/2}x_i^T\delta \le 0)$$
$$-\mathrm{pr}(\epsilon_i - n^{-1/2}x_i^T\delta' \le 0) + \mathrm{pr}(\epsilon_i - n^{-1/2}x_i^T\delta \le 0)\big\}\Big| = o_p(1),$$

where $||\cdot||$ denotes the L_2-norm. As a result,

$$
\begin{aligned}
& n^{-1/2}\Big|\sum_{i=1}^{n} x_{ij}\{I(\epsilon_i - n^{-1/2}x_i^T\widetilde{\delta} \le 0) - \tau\}\Big| \\
\le\ & n^{-1/2} \sup_{||\delta'-\delta||\le\Delta} \Big|\sum_{i=1}^{n} x_{ij}\{I(\epsilon_i - n^{-1/2}x_i^T\delta' \le 0) - I(\epsilon_i - n^{-1/2}x_i^T\delta \le 0) - \mathrm{pr}(\epsilon_i - n^{-1/2}x_i^T\delta' \le 0) \\
& +\mathrm{pr}(\epsilon_i - n^{-1/2}x_i^T\delta \le 0)\}\Big| + n^{-1/2} \sup_{||\delta'-\delta||\le\Delta} \Big|\sum_{i=1}^{n} x_{ij}\{\mathrm{pr}(\epsilon_i - n^{-1/2}x_i^T\delta' \le 0) - \mathrm{pr}(\epsilon_i - n^{-1/2}x_i^T\delta \le 0)\}\Big| \\
& +n^{-1/2}\Big|\sum_{i=1}^{n} x_{ij}\{I(\epsilon_i - n^{-1/2}x_i^T\delta \le 0) - \tau\}\Big| \\
=\ & o_p(1).
\end{aligned}
$$

Therefore, $\mathrm{pr}(j \in \widetilde{A}) \le \mathrm{pr}\{\xi_j(\widetilde{\delta}) = 0\} \to 0$, for $j = q+1, \ldots, p$. □

Proof of Lemma 2. Similarly as in the proof of Lemma 1, we can show that

$$
\begin{aligned}
& \sum_{i=1}^{n} \left\{\rho_\tau(\epsilon_i - x_i^T\delta/n^{1/2}) - \rho_\tau(\epsilon_i)\right\} + \lambda_n \sum_{j=1}^{p} \left(|\beta_{0j} + \delta_j/n^{1/2}| - |\beta_{0j}|\right) \\
\to\ & -\delta^T H + \delta^T B_1\delta/2 + \lambda_0 \sum_{j=1}^{p}\{|\delta_j| I(\beta_{0j} = 0) + \delta_j \mathrm{sign}(\beta_{0j}) I(\beta_{0j} \ne 0)\}
\end{aligned}
$$

in distribution. The result then follows from epi-convergence theory. □

Proof of Lemma A1. We have $V_{1n}^*(\delta) = n^{-1/2}\sum_{i=1}^n x_i^T\delta\big\{I(r_i|\hat{\epsilon}_i| < 0) - \tau\big\} = -n^{-1/2}\sum_{i=1}^n x_i^T\delta\big\{\tau - I(r_i < 0)\big\}$. Note that $E^*\{V_{1n}^*(\delta)\} = 0$ and $\mathrm{var}^*\{V_{1n}^*(\delta)\} = \tau(1-\tau)n^{-1}\sum_{i=1}^n \delta^T x_i x_i^T\delta \to \tau(1-\tau)\delta^T B_0\delta$ in probability. To check the Lindeberg condition, it suffices to show that $\forall\epsilon > 0$,

$$
n^{-1}\sum_{i=1}^{n} \mathrm{E}^*\Big(\Big[x_i^T\delta\big\{I(r_i|\hat{\epsilon}_i| < 0) - \tau\big\}\Big]^2 I\big[\big|x_i^T\delta\big\{I(r_i|\hat{\epsilon}_i| < 0) - \tau\big\}\big| > \epsilon\sqrt{n}\big]\Big) \to 0,
$$

in probability. This holds by noting that the left side of the above expression is upper bounded by $n^{-1}\sum_{i=1}^n (x_i^T\delta)^2 I(|x_i^T\delta| > \epsilon\sqrt{n})$, which converges to zero in probability by the dominated convergence theorem. The result of the lemma follows from the Lindeberg central limit theorem. □

Appendix 2

A Useful Lemma from Cheng & Huang (2010)

We use $r = \{r_1, \ldots, r_n\}$ to denote the random bootstrap weights and $z = \{z_1, \ldots, z_n\}$ to denote the random sample. Note that r and z induce two different sources of randomness. By the wild bootstrap mechanism, the distribution of r is independent of that of z. We adopt the following notation from Cheng & Huang (2010). A random quantity R_n is said to be $o_{p_r}^*(1)$ if for any $\epsilon, \delta > 0$, $\mathrm{pr}_z(\mathrm{pr}_{r|z}(|R_n| > \epsilon) > \delta) \to 0$, as $n \to \infty$. Similarly, R_n is said to be $O_{p_r}^*(1)$ in if for all $\delta > 0$ there exists a $0 < M < \infty$ such that $\mathrm{pr}_z(\mathrm{pr}_{r|z}(|R_n| > M) > \delta) \to 0$, as $n \to \infty$. And $o_{p_{r,z}}(1)$, $O_{p_{r,z}}(1)$ are the regular notion with respect to the joint probability distribution of r and z.

The following lemma from Cheng & Huang (2010) will be used repeatedly in our proof. It allows the transition of various stochastic orders in different probability spaces and leads to simplified proofs in many places.

LEMMA B1. *(Lemma 3 of Cheng & Huang (2010)) Suppose that*

$$Q_n = o^*_{p_r}(1), \quad R_n = O^*_{p_r}(1).$$

We have

$$\begin{aligned} A_n = o_{p_{r,z}}(1) &\Longleftrightarrow A_n = o^*_{p_r}(1), \\ B_n = O_{p_{r,z}}(1) &\Longleftrightarrow B_n = O^*_{p_r}(1), \\ C_n = Q_n \times O_{p_z}(1) &\Longleftrightarrow C_n = o^*_{p_r}(1), \\ D_n = R_n \times O_{p_z}(1) &\Longleftrightarrow D_n = O^*_{p_r}(1), \\ F_n = Q_n \times R_n &\Longleftrightarrow F_n = o^*_{p_r}(1). \end{aligned}$$

APPENDIX 3

Additional Examples of Random Weight Distribution

The random weights used in the wild residual bootstrap procedure are generated from a distribution G that satisfies Conditions 3–5 of the main paper. Two examples of such random weight distributions were given in Feng et al. (2011). We propose below three new weight distributions satisfying these conditions. Note that compared with the continuous distribution in Feng et al. (2011), the new distributions given in Examples 1–2 below have no restrictions on the value of τ.

Example 1.

$$\begin{aligned} g_1(r) = G_1'(r) = &-\frac{r}{8v_1}\mathrm{I}\left\{-2(\tau+v_1) \le r \le -2(\tau-v_1)\right\} \\ &+\frac{r}{8v_2}\mathrm{I}\left\{2(1-\tau-v_2) \le r \le 2(1-\tau+v_2)\right\}, \end{aligned}$$

where $0 < v_1 < \tau$ and $0 < v_2 < 1-\tau$.

Example 2.

$$\begin{aligned} g_2(r) = G_2'(r) = &-\frac{r}{32v_1}\mathrm{I}\left\{-4(a+v_1) < r < -4(a-v_1)\right\} \\ &-\frac{r}{32v_2}\mathrm{I}\left\{-4(\tau-a+v_2) < r < -4(\tau-a-v_2)\right\} \\ &+\frac{r}{32v_3}\mathrm{I}\left\{4(b-v_3) < r < 4(b+v_2)\right\} \\ &+\frac{r}{32v_4}\mathrm{I}\left\{4(1-\tau-b-v_3) < r < 4(1-\tau-b+v_2)\right\}, \end{aligned}$$

where $0 < v_1 < a$, $0 < v_2 < \tau - a$, $0 < v_3 < b$, $0 < v_4 < 1-\tau-b$, $0 < a < \tau$, and $0 < b < 1-\tau$.

Example 3. The point mass distribution

$$P(W = r) = a\mathrm{I}\{r = -4a\} + (\tau - a)\mathrm{I}\{r = -4(\tau - a)\} \\ + b\mathrm{I}\{r = 4b\} + (1 - \tau - b)\mathrm{I}\{r = 4(1 - \tau - b)\},$$

where $0 < a < \tau$ and $0 < b < 1 - \tau$.

Appendix 4

Additional Numerical Results

In Table 1, we summarize the simulation results on the comparison of empirical coverage probabilities ($\times 100$) and average interval lengths (in parentheses) for 95% confidence intervals for $\tau =0{\cdot}5$, $n = 250$ and $\tau =0{\cdot}7$, $n = 400$ for the various methods described in Section 4.1 of the main paper. We note that the standard errors of the coverage probabilities are below 0·01 and the standard errors of the confidence interval lengths are below 0·005 for all cases. These results supplement those in Table 1 of the main paper and demonstrate further improvement with increased sample size.

Figure 1 displays the QQ plots of the quantiles of the wild residual bootstrapped estimator versus the empirical quantiles of the corresponding penalized estimator for estimating the smallest coefficient $\beta_3 =0{\cdot}25$ for both the L_1 penalty and the adaptive L_1 penalty when sample size $n = 250$ and 400, for $\tau =0{\cdot}5$ and 0·7, respectively. Overall, the wild residual bootstrapped distribution has satisfactory performance.

Acknowledgments

Lan Wang and Adam Maidman acknowledge financial support from the U.S. National Science Foundation. Ingrid Van Keilegom acknowledges financial support from the European Research Council and a Interuniversity Attraction Pole research network of the Belgian government.

References

Boyd, S. & Vandenberghe, L. (2004). *Convex Optimization.* Cambridge: Cambridge University Press.
Cheng, G. & Huang, J. Z. (2010). Bootstrap consistency for general semiparametric M-estimation. *The Annals of Statistics* **38**, 2884–2915.
Feng, X., He, X. & Hu, J. (2011). Wild bootstrap for quantile regression. *Biometrika* **98**, 995–999.
Knight, K. (1998). Limiting distributions for l_1 regression estimators under general conditions. *The Annals of Statistics* **26**, 755–770.
Koenker, R. (2005). *Quantile Regression.* Cambridge: Cambridge University Press.
Wang, L., Wu, Y. & Li, R. (2012). Quantile regression for analyzing heterogeneity in ultra-high dimension. *Journal of the American Statistical Association* **107**, 214–222.
Wu, Y. & Liu, Y. (2009). Variable selection in quantile regression. *Statistica Sinica* **19**, 801–817.

Table 1: Empirical coverage probabilities ($\times 100$) and average interval lengths (in parentheses) for nominal 95% confidence intervals

	$\beta_1 = \Phi^{-1}(\tau)$	$\beta_3 = 0{\cdot}25$	$\beta_5 = 0{\cdot}5$	$\beta_7 = 1$	$\beta_9 = 2$	Zeros	TP	FP
			τ =0·5	$n = 250$				
New AL1	92·4 (0·22)	94·8 (0·09)	93·3 (0·09)	94·9 (0·07)	94·0 (0·08)	99·0 (0·03)	4	0·2
New AL2	91·4 (0·28)	94·3 (0·09)	93·5 (0·09)	93·9 (0·07)	93·8 (0·08)	99·2 (0·03)	4	0·2
New L1	93·6 (0·14)	93·9 (0·10)	93·0 (0·09)	95·3 (0·08)	94·3 (0·09)	95·1 (0·05)	4	2·8
New L2	92·6 (0·15)	94·4 (0·10)	92·9 (0·09)	95·4 (0·08)	94·5 (0·09)	95·0 (0·05)	4	2·8
Full RS	93·9 (0·37)	97·0 (0·12)	95·2 (0·11)	96·5 (0·09)	96·1 (0·11)	96·7 (0·10)	4	6
Full WB	91·5 (0·35)	97·0 (0·11)	96·2 (0·11)	97·2 (0·09)	96·2 (0·10)	96·7 (0·10)	4	6
TS AL RS	95·1 (0·34)	97·6 (0·14)	97·5 (0·11)	97·1 (0·13)	98·4 (0·15)	99·3 (0·11)	4	0·2
TS AL WB	93·2 (0·32)	93·6 (0·11)	95·8 (0·09)	95·4 (0·10)	97·2 (0·10)	99·1 (0·09)	4	0·2
TS L RS	94·7 (0·33)	96·7 (0·13)	96·8 (0·12)	97·5 (0·12)	97·8 (0·14)	96·9 (0·12)	4	2·8
TS L WB	93·2 (0·32)	93·8 (0·11)	94·8 (0·10)	96·0 (0·10)	96·3 (0·11)	96·6 (0·10)	4	2·8
Oracle RS	-	98·6 (0·14)	96·2 (0·12)	98·4 (0·11)	97·9 (0·13)	-	4	0
Oracle WB	-	96·4 (0·10)	95·0 (0·10)	96·1 (0·08)	95·7 (0·09)	-	4	0
			τ =0·7	$n = 400$				
New AL1	91·6 (0·27)	93·6 (0·07)	94·7 (0·06)	93·5 (0·07)	93·8 (0·07)	98·9 (0·03)	5	0·1
New AL2	91·6 (0·27)	93·8 (0·07)	94·8 (0·06)	93·6 (0·07)	94·3 (0·07)	99·2 (0·04)	5	0·0
New L1	92·5 (0·27)	93·8 (0·08)	95·2 (0·07)	94·7 (0·08)	93·9 (0·07)	95·6 (0·04)	5	2·1
New L2	92·8 (0·27)	94·0 (0·08)	95·6 (0·07)	94·6 (0·08)	94·0 (0·07)	96·1 (0·04)	5	2·1
Full RS	95·6 (0·30)	96·1 (0·09)	95·1 (0·08)	95·8 (0·09)	96·1 (0·08)	96·0 (0·08)	5	5
Full WB	93·1 (0·29)	95·6 (0·09)	96·1 (0·08)	95·2 (0·08)	95·4 (0·08)	95·9 (0·08)	5	5
TS AL RS	95·3 (0·30)	96·3 (0·07)	96·1 (0·08)	95·4 (0·08)	94·5 (0·08)	99·3 (0·09)	5	0·1
TS AL WB	92·7 (0·28)	96·6 (0·07)	96·9 (0·07)	96·1 (0·08)	95·5 (0·08)	99·2 (0·08)	5	0·1
TS L RS	94·8 (0·30)	96·0 (0·08)	95·4 (0·08)	95·5 (0·08)	95·0 (0·08)	95·5 (0·08)	5	2·1
TS L WB	92·4 (0·28)	96·4 (0·07)	96·7 (0·08)	95·2 (0·08)	95·6 (0·08)	96·0 (0·08)	5	2·1
Oracle RS	95·6 (0·30)	96·3 (0·08)	95·6 (0·07)	95·9 (0·08)	95·9 (0·08)	-	5	0
Oracle WB	92·8 (0·28)	95·8 (0·08)	96·8 (0·07)	95·7 (0·08)	95·4 (0·07)	-	5	0

New AL1: adaptive L_1 method with wild residual bootstrap ($\gamma = 1$); New AL2: adaptive L_1 method with wild residual bootstrap ($\gamma = 2$); New L1: L_1 method with modified wild residual bootstrap (data-driven choice of a_n); New L2: L_1 method with modified wild residual bootstrap ($a_n = n^{-1/3}$); Full RS: full model with rank-score method; Full WB: full model with wild residual bootstrap; TS AL RS: two-step procedure, adaptive L_1 ($\gamma = 1$) followed by rank-score method for the refitted model; TS AL WB: two-step procedure, adaptive L_1 ($\gamma = 1$) followed by wild residual bootstrap for the refitted model; TS L RS: two-step procedure, lasso followed by rank-score method for the refitted model; TS L WB: two-step procedure, lasso followed by wild residual bootstrap for the refitted model; Oracle RS: oracle model with rank-score method; Oracle WB: oracle model with wild residual bootstrap; Zeros: the reported average coverage probability (length) is the average for all zero coefficients; TP: average number of true positives; FP: average number of false positives.

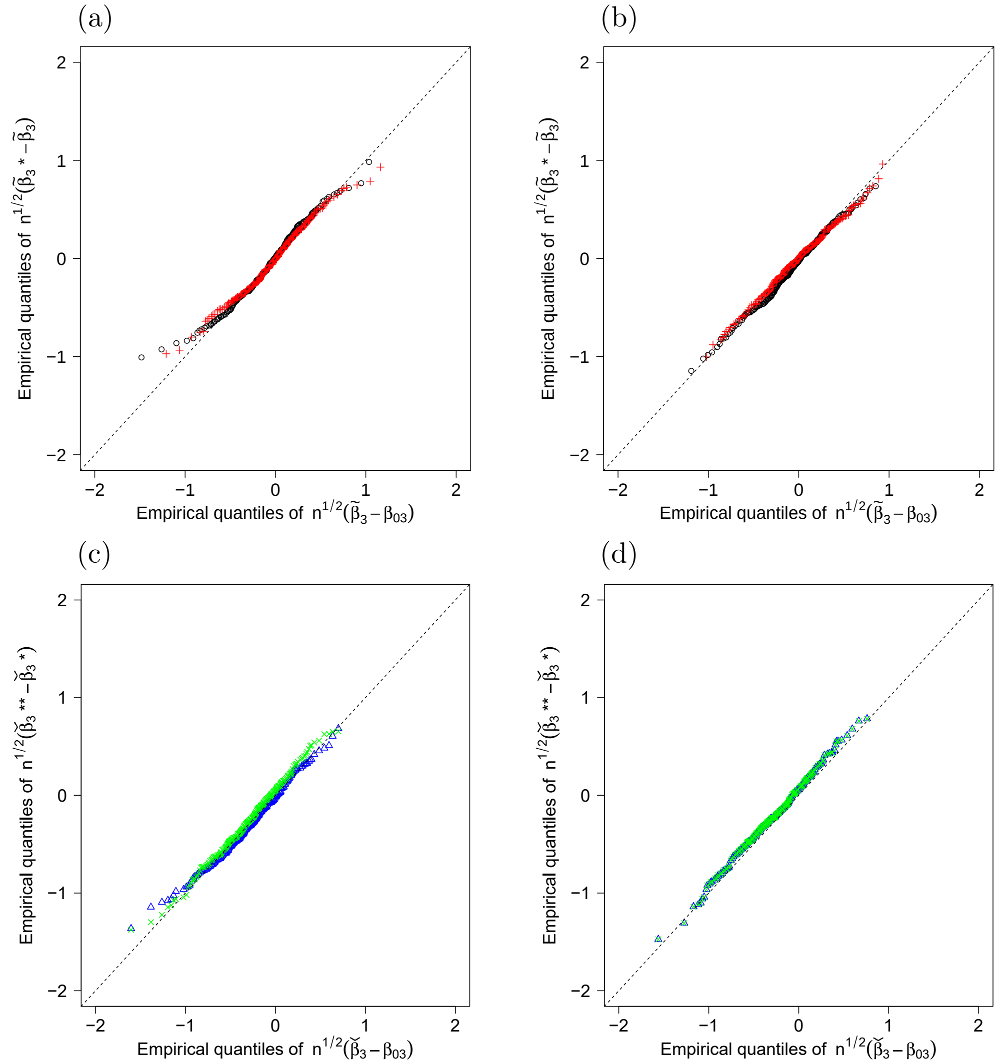

Fig. 1: QQ plots for the New AL1 (○), New AL2 (+), New L1 (△), and New L2 (×) methods for estimating $\beta_3 = 0{\cdot}25$ when $n = 250$. (a) and (b) adaptive L_1 method when $\tau = 0{\cdot}5$ and $\tau = 0{\cdot}7$, respectively; (c) and (d) L_1 method when $\tau = 0{\cdot}5$ and $\tau = 0{\cdot}7$, respectively.